National Computational Infrastructure for Lattice Gauge Theory

Objective

Create the software and hardware infrastructure needed for terascale simulations of quantum chromodynamics (QCD)

Executive Committee

R. Brower (BU), N. Christ (Columbia), M. Creutz (BNL), P. Mackenzie (FNAL), J. Negele (MIT),
C. Rebbi (BU), S. Sharpe (Washington), R. Sugar (UCSB) and W. Watson, III (JLab)

Scientific Goals

Non-perturbative study of QCD

- Calculation of weak decays of strongly interacting particles
  - * Determination of least well known parameters of the Standard Model
  - * Precision tests of the Standard Model
- Investigation of matter under extreme conditions
  - * Mapping the phase diagram of strongly interacting matter
  - * Determination of properties of the Quark Gluon Plasma
- Understanding the structure and interactions of hadrons
  - * Calculation of the spectrum of hadrons existing in nature and exploration of their interactions
  - * Determination of the quark and gluon structure of the nucleon and other hadrons

Project Goals

Create a unified programming environment that will enable the US lattice community to achieve very high efficiency on diverse multi-terascale hardware

- Portable, scalable software
- High-performance optimization on two target architectures
- Exploitation and optimization of existing application base
- Infrastructure to support national community
- Sharing of valuable lattice data and data management

Distributed Terascale Facility

The U.S. Lattice community five-year plan includes 3 machines of scale ten teraflops, sited at

- Brookhaven National Laboratory
- Fermi National Accelerator Laboratory
- Thomas Jefferson National Accelerator Facility

Impact on DOE Research

Lattice QCD calculations are essential to extract full physics potential from DOE's investment in frontier experimental facilities in nuclear and particle physics:

- Bates
- CEBAF
- Fermilab
- RHIC
- SLAC B-Factory

QCD API Design

Level 3

Dirac Operators, CG Routines, etc.

Level 2

Data Parallel QCD Lattice Wide Operations (overlapping Algebra and Messaging)	
Linear Algebra e.g. A = B * C	Data Movement SHIFT(A, mu)

Level 1

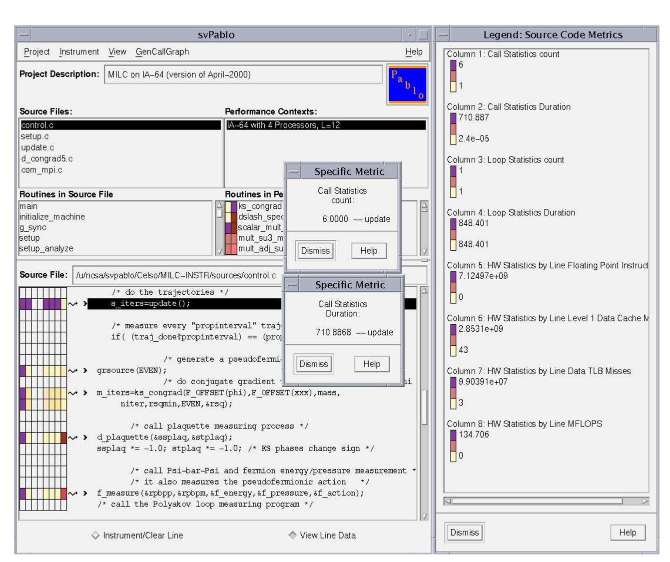
Single Node Lin Alg e.g. SU(3), Dirac	Message Passing Map Lat to Network

Data Grid

The Lattice Portal will give access to all QCD data, such as generated lattice configurations

- Replicated data (multi-site), global tree structured name space (like a local file system)
- Replica catalog, using an SQL database as a back end, replicated to multiple sites for fault tolerance
- Browse by attributes as well as by name
- Parallel file transfers (bbftp, gridftp, jparss,...)
- Drag-n-drop between sites (gui)
- Policy based replication (auto migrate between sites)

Application Optimization Tools



GUI interface to the Scalable Performance Tool *SvPablo*

QCD API Status and Schedule*

- Level 1 Linear Algebra 1st draft completed.
- Level 1 Message Passing (MP-API) design completed.

Implementation in MPI completed	Sept 2001
Myrinet optimization on GM	Feb 2002
C++ MP-API for QCDOC	begun
Application Port to MP-API	begun

- Demo of Level 2 "vertical slice", Feb 2002
- Optimization of Lin Alg for P4 by March 2002

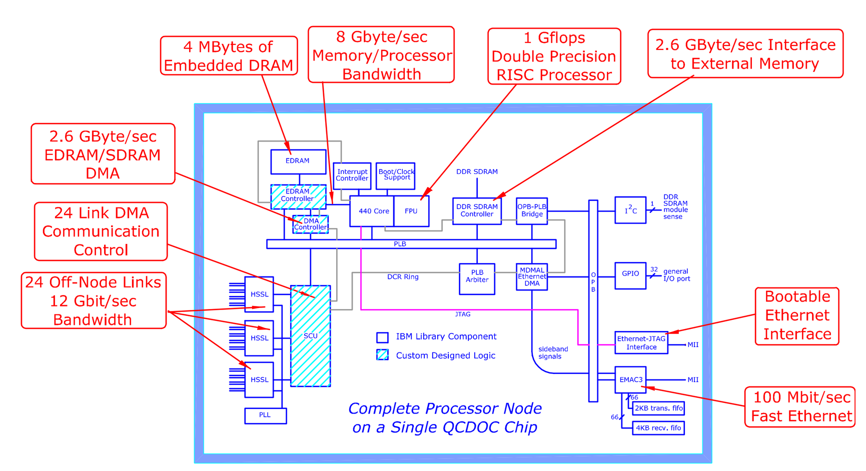
* `http://physics.bu.edu/~brower/SciDAC`

QCDOC Project

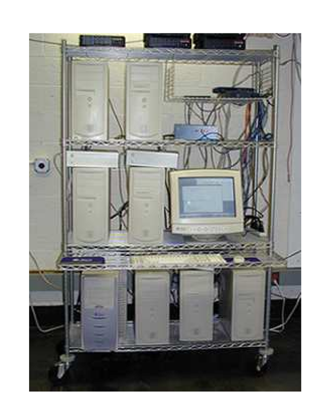
- Multi-teraflops QCD requirements scale as $\sim (N_{\text{lattice sites}})^{2.5}$ requiring:
  - * Thousands of processors.
  - * High-bandwidth, low-latency network.
- QCDOC (= QCD on a chip) architecture exploits lattice QCD regularity and supports $\sim 20K$ 1 Gflops processors.
- CPU/FPU and communications hardware integrated on a single chip manufactured by IBM.

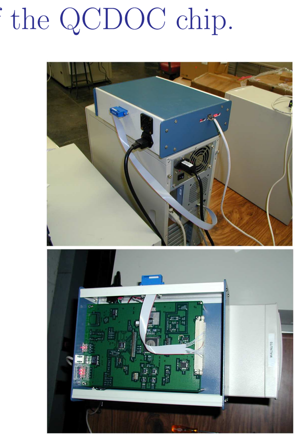
QCDOC Architecture

- Single-chip processing node:
  - * 1 Gflops, 64 bit IEEE FPU, PowerPC.
  - * 4 Mbytes on-chip memory.
- 6-dim mesh network.
  - * 1.5 Gbyte/sec/node network bandwidth.
  - * $\leq$ 500 ns network latency.
- $\leq$ 2 Gbyte DIMM DDR SDRAM external memory per node.
- Commodity Fast Ethernet provides independent host access to each node.



Block diagram of the QCDOC chip.





Four-node development system. Ethernet/JTAG prototypes.

QCDOC Status and Schedule

- ASIC design nearly finished.
- Physics code running in simulator:

single node FPU, SU(3)×2-spinor	84% (L1 cache) 78% (eDRAM)
cache fill/flush to eDRAM	3.2 GB/s sustained
bandwidth from/to eDRAM	2.5/2.0 GB/s
DDR fill/flush	1.2/1.7 GB/s

- 2-node Ethernet simulation.
- 4-node OS development system: PowerPC 405GP and Ethernet/JTAG boards.
- preliminary schedule:

  $\sim$ 1.5 TFlops sustained in 2002
  $\sim$ 10 TFlops sustained in 2003

Commodity Clusters

Flexible and powerful

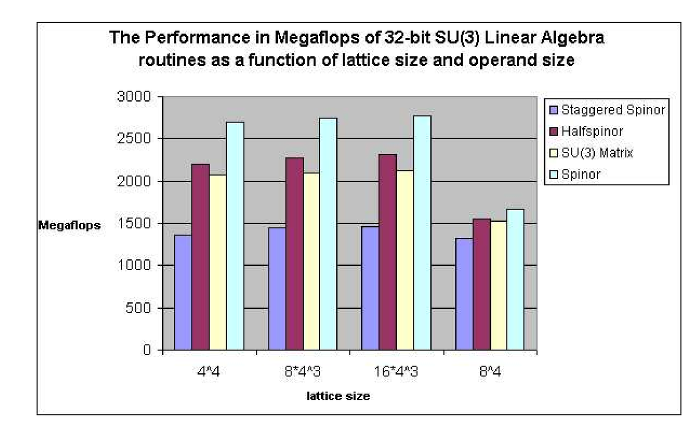
- Exploit commodity processor board and computer network engineering.
- Exploit commodity software (Linux OS, open source software, MPI, ...)
- Program, run legacy codes effortlessly.

Cost effective

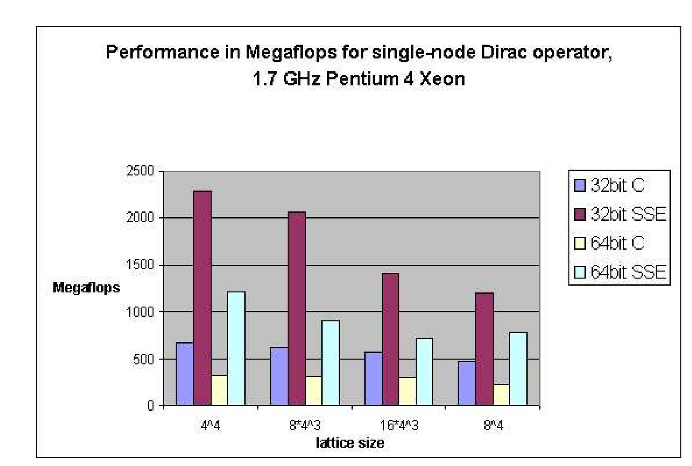
- Price/performance on lattice codes down to $1/MF on single node Pentium 4s.
- Clone, upgrade continuously and cheaply.
- Steady-state: continual upgrades, several thousand nodes, replace oldest third of system each year.

Performance on Pentium IV

Best price-performance for lattice QCD codes on commodity hardware currently provided by Pentium IV



Linear algebra on Pentium IV



Dirac operator on the Pentium IV

Exploitation of SSE instructions allows 2 Gflop/sec for the single-precision Dirac operator on a single processor, for problems residing in cache.

Cluster Status and Schedule

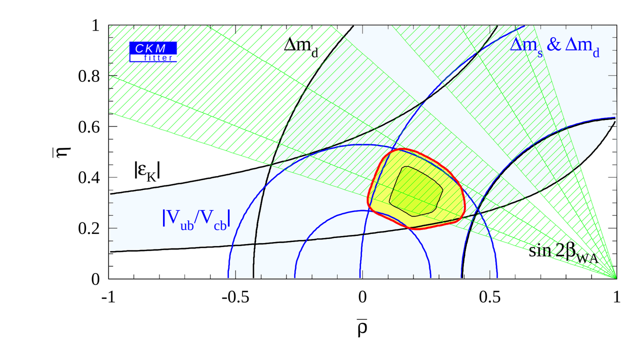
- Fermilab and JLAB/MIT designed and operated clusters for lattice QCD since 1999.
- Currently, 160 Pentium III processors (80 duals) at Fermilab, 88 ALPHA processors (singles, duals, quads) at JLAB and MIT; connected with Myrinet.
- Currently purchasing 200 node system at JLAB and 300 node system at FNAL, using dual Pentium IV and Myrinet 2000 interconnect.
- preliminary schedule:

  $\sim$ 1/4 TFlops sustained in 2002
  $\sim$ 10 TFlops sustained in 2005

Weak Decays

Some fundamental parameters of particle physics can be experimentally extracted only with aid of lattice gauge theory calculations.

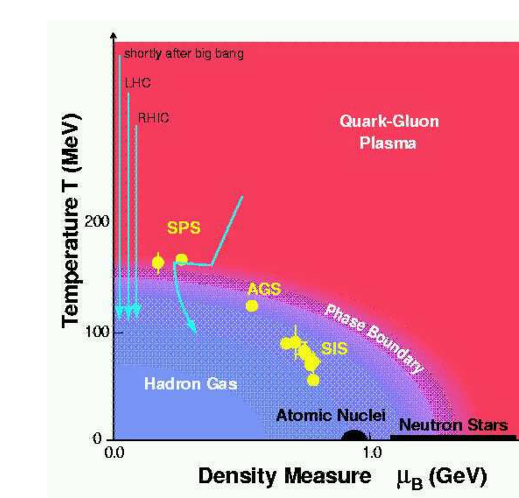
The constraints on the ρ and η quark transition parameters require both experimental measurements and accurate lattice calculations; a non-zero η leads to violation of CP symmetry as observed in Kaon decays and as needed to explain the matter-antimatter asymmetry of the universe. Any disagreements between the determinations signal a breakdown of the standard model of particle physics. Nearly all uncertainties will soon be dominated by lattice QCD uncertainties which will be very significantly reduced by the work of this project.



K. Anikeev et al., FERMILAB-Pub-01/197

The Quark-Gluon Plasma

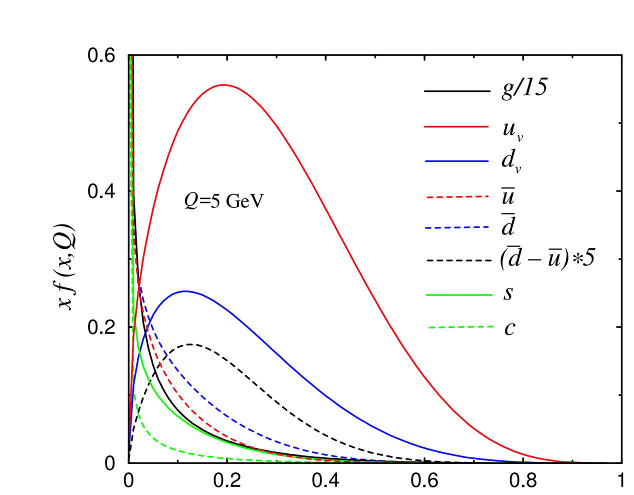
Last seen a few microseconds after the big bang, the quark gluon plasma is the quarry of the RHIC facility, and can be explored from first principles using lattice gauge theory.



QCD phase diagram
http://www-aix.gsi.de/ alice/phase-diag.jpg

Hadron Structure

High energy scattering experiments have measured the distribution of quarks and gluons in the proton. Ten teraflops sustained facilities will enable calculation of the moments of the distributions from first principles.



The quark and gluon distributions in the proton.